AD NUMBER
ADB103741
LIMITATION CHANGES

TO:

Approved for public release; distribution is unlimited.

FROM:

Distribution authorized to U.S. Gov't. agencies only; Administrative/Operational Use; 12 AUG 1986. Other requests shall be referred to Commandant of the Marine Corps, Attn: Code RD, Washington, DC 20380.
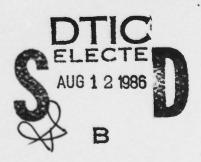
AUTHORITY

ONR/CNA ltr dtd 15 Dec 1988

AD-B103 741

RESEARCH MEMORANDUM

EFFECTS OF TRUNCATING A REFERENCE POPULATION ON CORRECTION OF VALIDITY COEFFICIENTS FOR RANGE RESTRICTION

Milton H. Maier

DTIC FILE COPY

DTIC
ELECTED
S
AUG 1 2 1986
B

CENTER FOR NAVAL ANALYSES

86 8 4 101

AD-B103741

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
	GOV'T 12 AUG 1986 ADM/OP USE
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	Distribution limited to U.S. DOD agencies only. Administrative information contained. Other requests for this document must be referred to the Commandant of the Marine Corps (Code RD).

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
CRM 85-40	

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Center for Naval Analyses	CNA	Commandant of the Marine Corps (Code RD)

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
4401 Ford Avenue Alexandria, Virginia 22302-0268	Headquarters, Marine Corps Washington, DC 20380

8a. NAME OF FUNDING / ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
Office of Naval Research	ONR	N00014-83-C-0725

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
800 North Quincy Street Arlington, Virginia 22217	65153M	C0031		

11. TITLE (Include Security Classification)

Effects of Truncating a Reference Population on Correction of Validity Coefficients for Range Restriction

12. PERSONAL AUTHOR(S)
Milton H. Maier

13a. TYPE OF REPORT	13b. TIME COVERED		14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Apr 1985	TO Jun 1985	June 1985	25

16. SUPPLEMENTARY NOTATION

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Aptitude tests, ASVAB (Armed Services Vocational Aptitude Battery), Coefficients, Correlation, Jobs, Performance, Personnel selection, Population, Recruits, Standards, Statistical samples, Truncation
05	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Correlation coefficients based on samples from occupational specialties that differ in qualification standards cannot be compared. The sample coefficients need to be put on the same metric by correcting them to a common reference population. The purpose of this analysis is to evaluate the effects of truncating the reference population on the correlation coefficients and on the intercorrelation of performance measures. Population-wide estimates were computed in the full population and in the truncated population with the bottom 10 percent deleted.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Lt. Col. G.W. Russell	(202) 694-3491	RDS-40

19 June 1985

MEMORANDUM FOR DISTRIBUTION LIST

Subj: Center for Naval Analyses Research Memorandum 85-40

Encl: (1) CRM 85-40, "Effects of Truncating a Reference Population
             on Correction of Validity Coefficients for Range
             Restriction," June 1985

     1. Enclosure (1) is forwarded as a matter of possible
interest.

     2. This Research Memorandum describes our analysis of the
effects of truncating an aptitude test reference population on
estimates of correlation coefficients corrected to that reference
population. The results are particularly applicable to current
discussion on this issue within the Joint-Service Job Performance
Measurement Working Group.

CHRISTOPHER JEHN
Director
Marine Corps Operations
  Analysis Group

Copy to:
  DC/S RD&S (2)

Distribution List:
Reverse page

Subj: Center for Naval Analyses Research Memorandum 85-40

DISTRIBUTION LIST

Defense Training Data and Analysis Center
Dr. D. J. Armor, Consultant
Army Research Institute
Navy Personnel Research & Development Center
HQ U.S. Marine Corps (Code MPI-20)
HQ USAF/MPXOA
Air Force Human Resources Laboratory
National Research Council
HQDA (DAPE-MPE-CS)
U.S. Coast Guard
OP-135L

JUNE
CRM 85-40 / ~~May~~ 1985

EFFECTS OF TRUNCATING A REFERENCE POPULATION ON CORRECTION OF VALIDITY COEFFICIENTS FOR RANGE RESTRICTION

Milton H. Maier

Marine Corps Operations Analysis Group

DTIC
ELECTE
S AUG 1 2 1986 D

B

QUALITY INSPECTED 4

Accession For		
NTIS GRA&I	□	✓
DTIC TAB	✓	
Unannounced	□	
Justification		
By		
Distribution/		
Availability Codes		
	Avail and/or	
Dist	Special	
B-3		

ABSTRACT

Correlation coefficients based on samples from occupational specialties that differ in qualification standards cannot be compared. The sample coefficients need to be put on the same metric by correcting them to a common reference population. The purpose of this analysis is to evaluate the effects of truncating the reference population on the correlation coefficients and on the intercorrelation of performance measures. Population-wide estimates were computed in the full population and in the truncated population with the bottom 10 percent deleted.

EXECUTIVE SUMMARY

INTRODUCTION

The goal of the Joint-Service Job Performance Measurement Project
is to link enlistment standards and job performance. Central to the
analysis for this project is the computation and interpretation of
correlation coefficients. The correlation between the Armed Services
Vocational Aptitude Battery (ASVAB) subtests and performance measures is
used to evaluate predictive validity of the ASVAB, and hence the justi-
fication for basing qualifying standards on the ASVAB.

Each service has its own set of qualifying ASVAB scores for
enlistment and for assigning recruits to occupational specialties.
Stated another way, the recruits in some services are more highly
selected than in others. Also, recruits in some occupational
specialties are more highly selected than in others. Electronics
technicians, for example, are more highly selected than automotive
mechanics. The different degrees of selection, arising from different
qualifying standards, complicate the computation and interpretation of
correlation coefficients.

As a rule, the effect of selecting people for an occupational
specialty, which of course includes selection into the service, is to
lower the correlation coefficients compared to the values that would
result if a representative sample from the total population of potential
recruits were assigned to the specialty. Other things being equal, the
more severe the selection, the lower the correlation coefficients in the
selected sample. An additional complication is that the distributions
of scores on some ASVAB subtests and performance measures are more
affected by the selection process than are others. The net result is
that observed correlation coefficients cannot be interpreted directly or
compared to each other. To facilitate comparison they should all be put
on a common basis by estimating what their values would be in a
reference population.

An ad hoc group from the Joint Services Job Performance Measurement
Working Group and the National Academy of Sciences Advisory Committee
was convened in the fall of 1984 to study the problem and make recom-
mendations. The group quickly agreed that the correlations should be
put on a common basis, which is sometimes called "correction for range
restriction," or obtaining "population-wide" estimates. The group also
agreed that all ASVAB subtests rather than a single test score, such as
the Armed Forces Qualification Test (AFQT) or an aptitude composite,
should be used simultaneously to obtain the population-wide estimates.
Technically, using all the subtests requires using the multivariate
model, whereas for a single test the univariate model suffices. (This
distinction becomes important when presenting the findings.) A final
agreement was that the proper base group for obtaining population-wide

estimates is the 1980 Youth Population, composed of the 18- through 23-year-old males and females in this country. The 1980 Youth Population was used to construct the 1980 ASVAB score scale introduced on 1 October 1984.

An unresolved point in correcting for range restriction is whether the full range of the reference population should be used or whether it should be truncated. The reason for truncating the population is to reduce the standard error of the population estimates. Other things being equal, the smaller the ratio of ASVAB subtest standard deviations in the population to those in the samples of selected recruits, the smaller the standard error of the population-wide estimates. Because standard errors are random variations that tend to obscure true values, they should be kept as low as feasible. A proposal to reduce the ratio of standard deviations and thereby the standard errors was to delete from the 1980 Youth Population the people who have AFQT scores below 10--in other words, to truncate the population at an AFQT score of 10.

The purpose of this analysis is to evaluate the effects of truncating the population on the validity coefficients of the ASVAB subtests and on the intercorrelation of the performance measures. Population-wide estimates were computed in the full population and in the truncated population with the bottom 10 percent deleted.

PROCEDURES

Two populations were considered. One was the World War II (WWII) Reference Population composed of males who served during WWII, and the other was the 1980 Youth Population. ASVAB subtest scores and performance measures were available for three Marine Corps occupational specialties--Ground Radio Repair, Automotive Mechanic, and Infantry Rifleman--used in the Marine Corps feasibility study on linking qualifying standards and job performance. The samples for two of the specialties, Radio Repair and Automotive Mechanic, had been tested with forms 6 and 7 of the ASVAB (ASVAB 6/7), which are on the WWII score scale. The Infantry Rifleman specialty had been tested with forms 8, 9, and 10 of the ASVAB (ASVAB 8/9/10), which are on the 1980 score scale. The standard deviations and correlation coefficients in the sample were corrected for range restriction, using both the full and truncated populations as the base.

FINDINGS

The findings germane to the purpose of evaluating the effects of truncating the population on validity coefficients and on the intercorrelation of the performance measures are as follows:

- The effects of deleting people with AFQT scores below 10 from the population are more complex in the multivariate model than indicated just by the ratio of standard

deviations. The ratio of covariances among the ASVAB subtests in the selected samples compared to the population has a greater effect on the population-wide estimates of the validity coefficients than does the ratio of standard deviations.

- There is no evidence in this analysis that the population-wide estimates based on correcting to a full population are more distorted by standard errors than those based on correcting to a truncated population.

Other findings of interest:

- The estimated validity coefficients of the ASVAB subtests showed less variability when corrected to the full population then when corrected to the truncated population. The implication of the lesser variability is that the statistical validity of decisions about classifying recruits into occupational specialties, such as distinguishing clerks from mechanics, appears to be less valid when the population estimates are based on the full population. Of course, the assignment decisions themselves are unaffected by the correction procedures.

- Measures of the spatial perception ability may be valid predictors of hands-on performance tests in some specialties.

RECOMMENDATIONS

- The Joint-Service Job Performance Measurement Working Group should adopt the recommendations of the ad hoc group:

  - Correlation coefficients should be corrected for range restriction.

  - The 1980 Youth Population (18- through 23-year-old males and females) should be the basis for correcting sample statistics.

  - All ASVAB subtests should be used as the explicit selection variables (use multivariate correlation model).

- The full-range 1980 Youth Population should be used as the basis for estimating population values.

TABLE OF CONTENTS

LIST OF TABLES

EFFECTS OF TRUNCATING A REFERENCE POPULATION ON CORRECTION
OF VALIDITY COEFFICIENTS FOR RANGE RESTRICTION

BACKGROUND

A major concern of the Joint-Service Job Performance Measurement
Working Group (JPMWG) is to develop and evaluate measures of job
performance. The job performance measures are administered to samples
of people who are working in the selected occupational specialties.
Before being assigned to an occupational specialty, military recruits
must obtain qualifying aptitude scores on the Armed Services Vocational
Aptitude Battery (ASVAB). Different occupational specialties have
different qualifying aptitude standards, which means the samples have
been subjected to different degrees of selection on the ASVAB.

The summary statistic most often used to show the degree of
relationship among performance and aptitude measures is the correlation
coefficient. Its value is affected by the extent to which the samples
have been selected on the basis of ASVAB scores. All coefficients of
the same value should reflect the same degree of relationship among the
variables. But, because the samples are subject to different degrees of
selection, the correlation coefficients are not directly comparable; an
adjustment is needed to put them on the same scale, or metric.

Members of the JPMWG and the National Academy of Sciences Advisory
Committee met in the fall of 1984 to consider procedures for computing
and reporting the degree of relationship among the variables. The
ad hoc group quickly agreed that the correlation coefficient is the most
useful summary statistic and that the coefficients should be corrected
for restriction in range. The correction procedure uses the regression
statistics computed in the samples to estimate the correlation that
would be obtained in the full population of all people who might have
been assigned to the specialty if there were no qualifying standards.
The population-wide estimates are on the same scale and directly
comparable to each other.

The ad hoc group recommended that the 1980 Youth Population,
composed of 18- through 23-year-old males and females, serve as the base
population. It also recommended that all ASVAB subtests be used in
computing the population-wide estimates. An unresolved question is
whether the full population should be used as the base or whether the
people with Armed Forces Qualification Test (AFQT) scores in the bottom
10 percent should be deleted. An AFQT percentile score of 10 was chosen
because the people in the bottom 10 percent are barred from the military
service. The purpose of this report is to compare the results of using
the full population and the truncated population, with the bottom
10 percent on AFQT deleted, for computing population-wide estimates.

PROBLEM

Three assumptions are involved in computing the population-wide estimates or as sometimes called "corrected correlation coefficients," (Gulliksen [1]):

- The regression weights are the same in the sample and population.

- The errors of prediction are the same in the sample and population.

- The partial correlations among the incidental variables (those not directly involved in selecting people for the occupational specialties) are the same in the sample and in the population.

As the degree of selection increases, which means that fewer people are qualified for the specialty, these assumptions become more tenuous. For example, if only the top quarter of the population qualifies for a specialty, the regression statistics in the sample are based on only a small portion of the total score distribution. A small error, say in estimating regression weights, may be greatly multiplied when the sample results are extended to the full population. Specifically, the standard error of the corrected correlation coefficients increases as the degree of selection increases. The degree of selection may be expressed as the ratio of standard deviations in the population and to those in the sample. Linn [2] reports that for a population coefficient of .5 that the standard error increases as follows:

Ratio of standard deviations (population/sample)	Standard error of population-wide estimate
1.0	.075
1.2	.086
1.4	.098
1.6	.110
1.8	.123
2.0	.135

The sample size is 100. A ratio of 1.0 means that the population and sample standard deviations are equal (no selection), and a ratio of 2.0 means that the population value is twice that of the sample (rather severe selection). The purpose of considering a truncation of the population is to reduce the ratio of standard deviations and thereby the standard errors.

Two other conditions should be met in evaluating the effects of truncating the reference population. One is that the corrected ASVAB

-2-

subtest validity coefficients should not be biased by the truncation. The second is that the procedures used by the JPMWG should also be used by the Joint-Service Selection and Classification Working Group, which, among other things, is concerned with validating the ASVAB.

The first condition, that the results for the ASVAB subtests should not be biased, becomes important because the scores on the ASVAB subtests that compose the AFQT would be more affected by the truncation than would scores on the remaining ASVAB subtests. The implication is that the corrected validity coefficients of the subtests in the AFQT would be relatively lower in the truncated population than in the full population.

The reason for desiring compatibility of procedures between the two working groups is that the two sets of validation results will be compared with each other. The validity coefficients should be on the same scale.

The results presented in this report bear on the ratios of the standard deviations and on the relative magnitude of the population-wide estimates for the ASVAB subtests based on corrections to the full and truncated populations.

PROCEDURES

Two sets of population values are available. One is for the 1980 Youth Population, composed of 18- through 23-year-old males and females, and the other is a simulation of the World War II Reference Population, composed of males who served during World War II (WWII). WWII population values are available for forms 6 and 7 of the ASVAB (ASVAB 6/7).[1] The 1980 Youth Population was tested with form 8 of the ASVAB, which is parallel to the current version of the ASVAB, forms 11, 12, and 13 (ASVAB 11/12/13). Standard deviations and intercorrelations were computed for each version of the ASVAB (forms 6 and 7 or 8, 9, and 10) in the full and truncated populations.

Sample statistics--including standard deviations, intercorrelation of ASVAB subtests, validity coefficients of ASVAB subtest, and intercorrelation of performance measures--were available for three samples. These samples were used in a Marine Corps study that evaluated the feasibility of setting ASVAB qualification standards against hands-on job performance tests [3]. The three samples comprised Marines

---

1. The simulated WWII Population values were computed for a sample of 2,025 applicants for enlistment tested in January and February of 1980 with ASVAB 8, ASVAB 6/7, and form 7A of the AFQT. The sample was used to scale ASVAB 8/9/10 to the WWII Reference Population, using AFQT 7A as the reference test. The sample was weighted by AFQT 7A to represent the WWII Population.

assigned to the Ground Radio Repair, Automotive Mechanic, and Infantry Rifleman specialties. The first two specialties were tested with ASVAB 6/7, and the sample statistics were corrected to the WWII Reference Population. The Infantry Rifleman sample was tested with ASVAB 8/9/10, and the sample statistics were corrected to the 1980 Youth Population. The population-wide estimates based on corrections to the full and truncated populations were computed for each sample.

RESULTS

Effects on Population-Wide Estimates

Table 1 shows the standard deviations, their differences, and the mean intercorrelation of the ASVAB subtests in the 1980 and WWII full and truncated populations. For the 1980 Youth Population, the largest differences in standard deviations were for the two subtests in the Verbal score (WK and PC). The WK standard deviation declined by almost one-fourth of the original value ($z = .248$). The mean intercorrelation for the two speeded subtests (NO and CS) showed the largest drop (.17 and .16, respectively). The standard deviations for the math subtests (AR and MK) and technical subtests (AS, MC, and EI) had little change. The mean intercorrelation for the two math subtests also showed little change.

The effects of truncating were less for the WWII population (part B of table 1) than for the 1980 Youth Population (part A). The statistics for the four interest measures were hardly affected by the truncation.

The intercorrelation matrices are shown in table 2 for the 1980 Youth Population and table 3 for the WWII Population. The coefficients for the full and truncated Reference Populations are shown in each table. Note that the correlation between AS and the two speeded tests approaches zero in the truncated 1980 Youth Population (the correlation between AS and CS is .04, table 2). The intercorrelation of the interest measures in the WWII population (table 3) is low. The clerical (CA) and mechanical (CM) interests are negatively correlated in the full (-.03) and truncated (-.06) WWII populations.

The effects of truncating the population on the validity of the ASVAB subtests are shown in table 4 for the Ground Radio Repair specialty ($N = 60$), in table 5 for the Automotive Mechanic specialty ($N = 131$), and in table 6 for the Infantry Rifleman specialty ($N = 53$). Results are shown for both the hands-on and written performance tests. ASVAB 6/7 had been administered to the Radio Repair and Automotive Mechanic samples, and ASVAB 8/9/10 to the Infantry Rifleman sample. The results for the first two samples can be compared because they are both referenced to the WWII Population, but not with the results for the Infantry Rifleman sample, which is referenced to the 1980 Youth Population.

TABLE 1

SUBTEST STANDARD DEVIATIONS AND INTERCORRELATIONS IN FULL AND TRUNCATED[a] REFERENCE POPULATIONS

Part A: 1980 Youth population

ASVAB 8/9/10

Subtest		Standard deviation			Mean intercorrelation		
Title	Symbol	Full	Truncated	z^b	Full	Truncated	Difference
Word Knowledge	WK[c]	7.710	5.797	.248	.66	.54	.12
Paragraph Comprehension	PC[c]	3.355	2.619	.219	.61	.47	.14
Arithmetic Reasoning	AR[c]	7.373	6.838	.073	.66	.58	.08
Numerical Operations	NO[c]	10.800	8.689	.195	.54	.37	.17
Coding Speed	CS	16.763	14.755	.120	.47	.31	.16
General Science	GS	5.010	4.374	.127	.67	.57	.10
Math Knowledge	MK	6.393	6.125	.042	.62	.55	.07
Auto/Shop Information	AS	5.550	5.248	.054	.51	.39	.12
Mechanical Comprehension	MC	5.349	5.081	.050	.59	.50	.09
Electronics Information	EI	4.236	3.890	.082	.61	.51	.10

TABLE 1 (Continued)

Part B: World War II population

ASVAB 6/7

Subtest		Standard deviation			Mean intercorrelation[d]		
Title	Symbol	Full	Truncated	z^b	Full	Truncated	Difference
Cognitive subtests							
Word Knowledge	WK[e]	7.0	6.2	.11	.64	.56	.08
Arithmetic Reasoning	AR[e]	4.7	4.3	.09	.65	.58	.07
Space Perception	SP[e]	4.2	3.8	.10	.52	.41	.09
Numerical Operations	NO	10.5	9.7	.08	.49	.40	.09
General Science	GS	4.3	4.1	.05	.64	.59	.05
Math Knowledge	MK	4.9	4.7	.04	.60	.54	.06
Auto Information	AI	4.8	4.6	.04	.58	.51	.07
Shop Information	SI	4.2	3.8	.10	.60	.52	.08
Mechanical Comprehension	MC	4.5	4.2	.09	.65	.59	.06
Electronics Information	EI	5.7	5.2	.09	.64	.57	.07
Interest measures							
Mechanical	CM	4.2	4.2	.00	.16	.15	.01
Electronics	CE	4.6	4.6	.00	.20	.19	.01
Clerical	CA	3.0	3.0	.00	.20	.17	.03
Combat	CC	3.8	3.6	.05	.31	.28	.03

a. Bottom 10 percent of AFQT scores deleted.
b. Difference between standard deviations divided by full standard deviation.
c. Subtests in AFQT for ASVAB 8/9/10.
d. Means are based on cognitive subtests only.
e. Subtests in AFQT for ASVAB 6/7.

-6-

TABLE 2

INTERCORRELATION[a] OF ASVAB 8 IN 1980 YOUTH POPULATION

Part A: Full range

Subtest[c]

	PC	AR	NO[d]	CS[d]	GS	MK	AS	MC	EI
WK	80	71	62	55	80	67	53	60	68
PC		67	61	56	69	64	42	52	57
AR			63	52	72	83	53	69	66
NO				70	53	62	31	41	42
CS					45	52	23	34	34
GS						70	64	70	76
MK							42	60	59
AS								74	75
MC									75
EI									

Part B: Truncated[b]

Subtest[c]

	PC	AR	NO[d]	CS[d]	GS	MK	AS	MC	EI
WK	68	63	38	35	73	61	40	50	59
PC		57	39	39	56	56	26	39	42
AR			52	40	66	80	45	63	58
NO				60	32	54	10	24	22
CS					27	42	04	18	15
GS						64	56	64	71
MK							32	54	51
AS								71	71
MC									70
EI									

a. Decimals omitted.
b. Bottom 10 percent of AFQT scores (WK + PC + AR + NO/2) deleted from population.
c. See table 1 for subtest titles.
d. Scores adjusted for using military testing materials.

-7-

TABLE 3

ESTIMATED INTERCORRELATION[a] OF ASVAB 6/7 IN WORLD WAR II POPULATION

Part A: Full range

Subtest[c]

	WK	AR	SP	NO	CS	MK	AI	SI	MC	EI	CM	CE	CA	CC
WK		73	50	50	80	66	59	64	66	72	06	14	28	33
AR	73		58	63	69	78	55	57	69	64	07	21	28	28
SP	50	58		42	50	52	46	51	62	53	14	21	18	26
NO	50	63	42		49	63	40	43	48	45	03	18	25	24
CS	80	69	50	49		67	61	64	70	70	10	21	24	33
MK	66	78	52	63	67		47	47	63	59	02	24	27	23
AI	59	55	46	40	61	47		73	67	71	39	18	11	33
SI	64	57	51	43	64	47	73		67	71	30	14	09	39
MC	66	69	62	48	70	63	67	67		71	23	22	15	34
EI	72	64	53	45	70	59	71	71	71		26	26	18	32
CM	06	07	14	03	10	02	39	30	23	26		41	-03	26
CE	14	21	21	18	21	24	18	14	22	26	41		34	10
CA	28	28	18	25	24	27	11	09	15	18	-03	34		09
CC	33	28	26	24	33	23	33	39	34	32	26	10	09	

Part B: Truncated on AFQT[b]

Subtest[c]

	WK	AR	SP	NO	CS	MK	AI	SI	MC	EI	CM	CE	CA	CC
WK		65	36	39	76	60	52	55	59	65	03	14	25	25
AR	65		48	57	64	76	47	47	63	55	04	21	24	20
SP	36	48		31	40	43	36	40	55	42	12	21	13	17
NO	39	57	31		42	59	30	30	40	34	00	18	21	15
CS	76	64	40	42		63	55	58	66	66	07	21	20	27
MK	60	76	43	59	63		39	39	57	53	-02	24	24	16
AI	52	47	36	30	55	39		70	63	67	40	17	06	27
SI	55	47	40	30	58	39	70		62	64	32	13	02	32
MC	59	63	55	40	66	57	63	62		66	23	23	10	27
EI	65	55	42	34	66	53	67	64	66		26	27	13	25
CM	03	04	12	00	07	-02	40	32	23	26		39	-06	25
CE	14	21	21	18	21	24	17	13	23	27	39		35	09
CA	25	24	13	21	20	24	06	02	10	13	-06	35		03
CC	25	20	17	15	27	16	27	32	27	25	25	09	03	

a. Decimals omitted.
b. Estimated bottom 10 percent of AFQT scores (WK + AR + SP) deleted from population.
c. See table 1 for titles of subtests.

-8-

TABLE 4

EFFECTS OF TRUNCATED REFERENCE POPULATION ON CORRECTED VALIDITY COEFFICIENTS - GROUND RADIO REPAIRERS

ASVAB 6/7 subtest

		Standard deviation					Performance measure									
			Reference population		Ratio		Hands-on test					Written test				
							Validity			Difference		Validity			Difference	
Title	Symbol	Sample	Full	Truncated	Full	Truncated	Sample[a]	Full[b]	Truncated[c]	Full	Truncated	Sample[a]	Full[b]	Truncated[c]	Full	Truncated
Word Knowledge	WK	4.1	7.0	6.2	1.7	1.5	.27	.66	.58	.39	.31	.38	.70	.65	.32	.27
Arithmetic Reasoning	AR	2.4	4.7	4.3	2.0	1.8	.31	.68	.63	.37	.32	.22	.66	.61	.44	.39
Space Perception	SP	3.3	4.2	3.8	1.3	1.2	.40	.66	.59	.26	.19	-.04	.33	.21	.37	.25
Numerical Operations	NO	6.9	10.5	9.7	1.5	1.4	-.09	.41	.32	.50	.41	.23	.55	.49	.32	.26
General Science	GS	2.8	4.3	4.1	1.5	1.5	.17	.61	.56	.44	.39	.41	.69	.66	.28	.25
Math Knowledge	MK	2.4	4.9	4.1	2.0	1.7	.32	.63	.58	.31	.26	.24	.65	.62	.41	.38
Auto Information	AI	4.4	4.8	4.6	1.1	1.0	.25	.53	.45	.28	.20	.11	.46	.39	.35	.28
Shop Information	SI	3.1	4.2	3.8	1.4	1.2	.04	.44	.32	.40	.28	.24	.62	.55	.38	.31
Mechanical Comprehension	MC	3.0	4.5	4.2	1.5	1.4	.25	.54	.46	.29	.21	.20	.60	.55	.40	.35
Electronics Information	EI	3.8	5.7	5.2	1.5	1.4	.17	.59	.51	.42	.34	.24	.63	.57	.39	.33
Mechanical Interest	CM	4.7	4.2	4.2	0.9	0.9	-.15	-.18	-.22	-.03	-.07	.10	.11	.09	.01	-.01
Electronics Interest	CE	4.2	4.6	4.6	1.1	1.1	-.22	-.01	-.02	.21	.20	.11	.18	.18	.07	.07
Clerical Interest	CA	2.4	3.0	3.0	1.3	1.3	-.04	.28	.25	.32	.29	.28	.24	.21	-.04	-.07
Combat Interest	CC	4.3	3.8	3.6	0.9	0.8	.00	.26	.18	.26	.18	-.16	.16	.08	.32	.24

a. Coefficients computed in sample selected on basis of ASVAB scores.
b. Coefficients corrected to full reference population.
c. Coefficients corrected to truncated reference population.

-9-

TABLE 5

EFFECTS OF TRUNCATED REFERENCE POPULATION ON VALIDITY COEFFICIENTS - AUTOMOTIVE MECHANICS

ASVAB 6/7 subtest		Standard deviation					Performance measure									
			Reference population		Ratio		Hands-on test					Written test				
							Validity			Difference		Validity			Difference	
Title	Symbol	Sample	Full	Truncated	Full	Truncated	Sample[a]	Full[b]	Truncated[c]	Full	Truncated	Sample[a]	Full[b]	Truncated[c]	Full	Truncated
Word Knowledge	WK	5.0	7.0	6.2	1.4	1.2	.15	.37	.32	.22	.17	.35	.59	.53	.24	.18
Arithmetic Reasoning	AR	3.4	4.7	4.3	1.4	1.3	.11	.34	.28	.23	.17	.19	.50	.44	.31	.25
Space Perception	SP	3.6	4.2	3.8	1.2	1.1	.13	.32	.25	.19	.12	.20	.45	.36	.25	.16
Numerical Operations	NO	8.6	10.5	9.7	1.2	1.1	.09	.28	.22	.19	.13	.08	.35	.27	.27	.19
General Science	GS	3.3	4.3	4.1	1.3	1.2	.12	.36	.32	.24	.20	.33	.57	.53	.24	.20
Math Knowledge	MK	3.9	4.9	4.7	1.3	1.2	.10	.32	.27	.22	.17	.15	.47	.42	.32	.27
Auto Information	AI	4.8	4.8	4.6	1.0	1.0	.50	.57	.52	.07	.02	.41	.55	.50	.14	.09
Shop Information	SI	3.8	4.2	3.8	1.1	1.0	.40	.49	.46	.09	.06	.27	.46	.40	.19	.13
Mechanical Comprehension	NC	3.4	4.5	4.2	1.3	1.2	.33	.45	.42	.12	.09	.34	.55	.50	.21	.16
Electronics Information	EI	4.4	5.7	5.2	1.3	1.2	.30	.45	.42	.15	.12	.35	.57	.52	.22	.17
Mechanical Interest	CN	3.8	4.2	4.2	1.1	1.1	.31	.21	.21	-.10	-.10	.23	.13	.12	-.10	-.11
Electronics Interest	CE	4.1	4.6	4.6	1.1	1.1	.01	.22	.22	.21	.21	-.09	.11	.11	.20	.20
Clerical Interest	CA	2.9	3.0	3.0	1.0	1.0	-.14	.19	.17	.33	.31	-.18	.11	.07	.29	.25
Combat Interest	CC	3.9	3.8	3.6	1.0	0.9	-.08	.01	-.04	.09	.04	.00	.14	.08	.14	.08

a. Coefficients computed in sample selected on basis of ASVAB scores.
b. Coefficients corrected to full reference population.
c. Coefficients corrected to truncated reference population.

-10-

TABLE 6

EFFECTS OF TRUNCATED REFERENCE POPULATION ON VALIDITY COEFFICIENTS - INFANTRY RIFLEMEN

ASVAB 8/9/10 subtest

		Standard deviation					Performance measure										
		Reference population			Ratio		Hands-on test					Written test				
							Validity			Difference		Validity			Difference	
Title	Symbol	Sample	Full	Truncated	Full	Truncated	Sample[a]	Full[b]	Truncated[c]	Full	Truncated	Sample[a]	Full[b]	Truncated[c]	Full	Truncated
Word Knowledge	WK	5.0	7.7	5.8	1.5	1.2	.39	.53	.37	.14	-.02	.63	.82	.73	.19	.10
Paragraph Comprehension	PC	2.4	3.4	2.6	1.4	1.1	.48	.61	.49	.13	.01	.43	.69	.54	.26	.11
Arithmetic Reasoning	AR	5.1	7.4	6.8	1.4	1.3	.26	.43	.31	.17	.05	.44	.66	.57	.22	.13
Numerical Operations	NO	8.0	10.8	8.7	1.4	1.1	.14	.50	.34	.36	.20	.22	.61	.42	.39	.20
Coding Speed	CS	11.5	16.8	14.8	1.5	1.3	.13	.39	.23	.26	.10	.30	.57	.42	.27	.12
General Science	GS	3.9	5.0	4.4	1.3	1.1	.42	.55	.43	.13	.01	.56	.72	.62	.16	.06
Math Knowledge	MK	5.0	6.4	6.1	1.3	1.2	.45	.53	.44	.08	-.01	.51	.65	.58	.14	.07
Auto/Shop Information	AS	4.6	5.6	5.2	1.2	1.1	.42	.45	.35	.03	-.07	.28	.38	.23	.10	-.05
Mechanical Comprehension	MC	4.3	5.3	5.1	1.2	1.2	.50	.58	.50	.08	-.00	.41	.52	.40	.11	-.01
Electronics Information	EI	3.2	4.2	3.9	1.3	1.2	.45	.52	.41	.07	-.04	.47	.66	.56	.19	.09

a. Coefficients computed in sample selected on basis of ASVAB scores.

b. Coefficients corrected to full reference population.

c. Coefficients corrected to truncated reference population.

-11-

The Radio Repair sample was the most selected. The two math subtests plus GS and EI composed the Electronics Repair aptitude composite used to assign recruits to the Ground Radio Repair specialty. For the two math subtests the ratio of standard deviations between the full population and the sample for the two math subtests was 2.0. The ratios for the math subtests were about 10 percent lower between the truncated population and the sample.

For the radio repairers, the validity coefficients of the ASVAB subtests in both the full and truncated Reference Populations increased substantially compared to the sample values. The validity coefficients of the cognitive subtests in the truncated population were uniformly lower than in the full population. There was almost no shift in their rank order. In fact, the rank order of the cognitive subtests in the sample was about the same in both populations. Similar rank ordering of the subtests are obtained for both the hands-on and written tests.

The validity coefficients for the cognitive subtests varied more in the truncated than in the full population. The validity of the cognitive subtests was uniformly higher for predicting the written test score than for predicting hands-on test scores; a notable exception was the Space Perception (SP) subtest. SP was a highly respectable predictor of hands-on test scores (.66 in the full population), a poor predictor of the written test scores (.33), and a modest predictor of training grades (.44).

The results for the Automotive Mechanic sample (table 5) were similar to those for the Radio Repair sample. The selection effects were less, as shown by the smaller ratios of the standard deviations, and the corrected validity coefficients tended to be lower. The differences between the full and truncated population values tended to be about the same. The rank-order of the cognitive subtests was almost identical in the full and truncated populations. But, strangely the validity of the mechanical interest measure (CM) as a predictor of scores on both the hands-on and written performance tests was higher in the sample than in the populations. The reason lies in the correlation of CM with the cognitive subtests. For example, in the Mechanics sample, CM correlated .53 with AI, but in the full population the correlation was only .39.

The Rifleman sample had been tested with ASVAB 8/9/10, and the sample values were corrected to the 1980 Youth Population (table 6). The mean aptitude of the sample on the subtests was about one-third to one-half of a standard deviation above the mean of the population. The ratios of the standard deviations ranged from 1.5 for WK and CS to 1.2 for AS and MC. The ratios in the truncated population were lower and, of course, followed the pattern for the subtests shown earlier in table 1, part A. The magnitude of estimated validity coefficients in the full and truncated populations showed differential effects for the subtests. All the ratios of standard deviations were greater than 1.0,

yet four of the estimated validity coefficients against the hands-on test in the truncated population actually declined (WK, MK, AS, and EI), one remained constant (MC), and one increased by .20 (NO). The rest showed a modest increase. In the full population, the estimated validity increased for all subtests. Note that ASVAB 8/9/10 did not contain interest measures. Against the written performance test, only AS and MC showed a decline of estimated validity in the truncated population. Apparently the differences in patterns of covariances between the sample of riflemen and the truncated population were enough to lower the corrected validity coefficients.

Effects of Truncating the Population on the Intercorrelation of Performance Measures

The preceding results focused on the validity of the ASVAB subtests. These subtests figure directly in the selection process and are termed "explicit selection variables." In this subsection the focus is on the intercorrelation among the performance measures, which are affected incidentally by the selection process; that is, their variance and covariance are affected only to the extent that they correlate with the explicit selection variables. Variables of this type are said to be subject to "incidental selection" and are called "incidental variables."

The intercorrelation of these performance measures--hands-on tests, written tests, and training grades--for the three samples are shown in table 7. The degree of change in the population-wide estimates in each sample for these incidental variables corresponds to those found above for the ASVAB subtests. The largest change is for the Radio Repair sample, and the smallest is for the Rifleman sample in the truncated population. The pattern of intercorrelations in each sample shows little change between that for the sample and that for the corrected values in either the full or truncated populations.

DISCUSSION

The primary impetus for the analysis in this report arose from a concern to reduce the standard errors of the estimated population-wide correlation coefficients. In the univariate model, in which there is only one explicit selection variable, this standard error is a direct function of the ratio of standard deviations. In this analysis, however, the multivariate model was used. The multivariate model involves the ratio of variance-covariance matrices in the sample to that in the population. The effects on the population-wide estimates are therefore more complex than in the univariate model.

The pattern of corrected coefficients, of both the ASVAB subtests as explicit selection variables and the performance measures as incidental selection variables, is similar in the full and truncated populations. The corrected correlation coefficients are higher in the

TABLE 7

EFFECTS OF TRUNCATED REFERENCE POPULATION ON THE INTERCORRELATION OF PERFORMANCE MEASURES

Part A: Ground Radio Repair specialty

Performance measure	Sample		Reference population			
			Full[a]		Truncated[b]	
	HO	WR	HO	WR	HO	WR
Hands-on (HO)						
Written (WR)	.12		.48		.41	
Grades (GR)	.24	.31	.52	.62	.45	.57

Part B: Automotive Mechanic specialty

Performance measure	Sample		Reference population			
			Full[a]		Truncated[b]	
	HO	WR	HO	WR	HO	WR
Hands-on (HO)						
Written (GR)	.35		.45		.42	
Grades (GR)	.41	.55	.51	.69	.49	.65

Part C: Infantry Rifleman specialty

Performance measure	Sample		Reference population			
			Full[a]		Truncated[b]	
	HO	WR	HO	WR	HO	WR
Hands-on (HO)						
Written (WR)	.45		.58		.46	
Grades (GR)	.34	.54	.39	.61	.34	.59

---

a. Sample correlation coefficients corrected to full reference population.
b. Sample correlation coefficients corrected to truncated reference population.

full population, but the rank order of the coefficients remains approximately the same. There is no sign in these results that the estimated correlation coefficients in the full population are distorted by standard errors. The standard errors may be larger in the full population than in the truncated population, but the interpretation of the results would be similar for both sets of correlations.

The main conclusion from this analysis is that no apparent error is introduced by using the full population to correct the sample statistics. The full population has a clear definition--all 18- through 23-year-old males and females in this country--in contrast to the truncated population, which would always need to be footnoted. The full population has already been extensively used, notably to construct the 1980 ASVAB score scale introduced on 1 October 1984. The evidence is on the side of using the variance-covariance matrix of the full population as the basis for correcting the sample values for range restriction.

The analysis produced other findings that were not directly germane to the issue of choosing the appropriate base population. These findings are discussed below.

Validity Generalization

The variability among the ASVAB subtest validity coefficients is related to whether are corrected to the full or truncated population. The standard deviations of the validity coefficients are:

| | Performance measure | | | | | |
| | Hands-on test | | | Written test | | |
Specialty	Sample	Full	Truncated	Sample	Full	Truncated
Radio Repair	.144	.093	.111	.126	.115	.138
Automotive Mechanic	.147	.091	.100	.107	.075	.086
Infantry Rifleman	.137	.068	.084	.129	.120	.141

The variability among the validity coefficients of the ASVAB subtests is larger in the truncated population than in the full population. The apparent differential validity of the ASVAB--that is, the validity coefficients for the ASVAB subtests are different for different occupational specialties--could be improved by using the truncated population variance-covariance matrix as the base.

The sets of validity coefficients corrected to the full population lend more support to the validity-generalization argument that all cognitive tests tend to be valid for all occupations. In fact, with the exception of the speeded subtests (NO and CS), SP, and AS, the estimated subtest validity coefficients in the full populations are similar for each specialty.

The question of differential validity is crucial to using ASVAB for assigning recruits to different occupational specialties. The similar patterns of validity coefficients across the three specialties examined indicate that the differential validity of the ASVAB is modest. Improvements would be best obtained by developing new predictors to measure aptitudes not currently covered by the ASVAB rather than by truncating the population.

Validity of the Space Perception Subtest

Measures of spatial perception traditionally have been included in multiple aptitude batteries. Forms 6 and 7 not only contained SP, but it was part of the AFQT. It was dropped, however, when forms 8, 9, and 10 were developed. One reason is that females as a group score lower on spatial perception than males. Another reason is that SP was found to have little unique validity against the traditional criterion measure of final grades in occupational specialty training courses. The estimated validity of SP in the full population for predicting scores on the three performance measures is as follows (for comparison, the validity of AR is shown in parentheses[1]):

	Performance measure		
Specialty	Hands-on	Written	Grades
Radio Repair	.66 (.68)	.33 (.66)	.44 (.71)
Automotive Mechanic	.32 (.34)	.45 (.50)	.49 (.66)
Infantry Rifleman[2]	.50 (.66)	.47 (.58)	–

Because SP is more independent of the other ASVAB subtests than is AR, its unique validity for predicting hands-on test scores is relatively higher; that is, in a multiple regression equation, SP would have relatively higher beta weights than AR when predicting hands-on performance measures than when predicting written tests or training grades. The suggestion is that SP may be a valid predictor of hands-on test scores, and hence may have a legitimate place in the ASVAB.

Speeded Tests

What the speeded tests measure appears to depend in part on the group being tested. CS and NO are frequently called tests of "perceptual speed and accuracy." For most of the population that may be an accurate label. The measures of perceptual speed and accuracy are

---

1. AR was chosen because of its high mean intercorrelation with other ASVAB subtests and its high mean validity across occupational specialties.
2. Sample tested with ASVAB 6/7, N = 140; training grades not available for full sample.

relatively independent of the other ASVAB subtests, as shown by the low mean intercorrelations (table 1). The intercorrelations suggest, however, that for people with low aptitude the speeded subtests have a noticeable cognitive component. The mean intercorrelation of NO and CS showed a large drop in the 1980 Youth Population when the bottom 10 percent on AFQT was deleted (from .54 to .37 for NO, and .47 to .31 for CS). Other analyses of the 1980 Youth Population show that NO and CS are more highly intercorrelated with the other subtests for groups that have low mean aptitude (e.g. non-high school graduates from racial or ethnic minorities).

Effects on Math and Verbal Subtests

Truncating the 1980 Youth Population affected the verbal subtests (WK, PC, and GS) and NO more than the math subtests (AR and MK). The reason is that the math subtests have relatively few easy items. The verbal subtests and NO have many easy items, which spread out the people who score at the low end of the scale. The minimum raw scores for the verbal tests and NO are more than three standard deviations below the mean; the minimum subtest standard scores for these subtests are truncated at 20, three standard deviations below the mean (standard deviation equals 10). The minimum AR and MK raw scores are less than three standard deviations below the mean (standard scores of 26 for AR and 29 for MK). The discriminations at the low end of the AFQT scale therefore are primarily a function of WK, PC, and NO, rather than of AR.

The discussion of standard scores raises one more point about the appropriate variance-covariance matrix. The analysis in this report, like other analyses that estimated population values, used ASVAB subtest raw scores, rather than subtest standard scores, to compute the population variance-covariance matrix. Because WK, PC, GS, and NO are truncated when computing standard scores, the population variance-covariance matrix for subtest standard scores has slightly different values than the one using raw scores. Subtest standard scores are used in the operational testing program, and they of course should be used to compute the population variance-covariance matrix. The appropriate matrix will be presented in a forthcoming CNA report on the 1980 score scale.

RECOMMENDATIONS

- The Joint-Service Job Performance Measurement Working Group should adopt the recommendations of the ad hoc group:

    - Correlation coefficients should be corrected for range restriction.

- The 1980 Youth Population (18- through 23-year-old males and females) should be the basis for correcting sample statistics.

- All ASVAB subtests should be used as the explicit selection variables (use multivariate model).

● The full-range 1980 Youth Population should be used as the basis for estimating population values.

REFERENCES

[1] Gulliksen, Harold. Theory of Mental Tests. New York: Wiley, 1950

[2] Linn, R. L., "Corrections for Range Restriction (Draft)," prepared
     for the National Academy of Sciences Committee on Job Performance
     Measurement, Unclassified, 10 Aug 1984

[3] CNA, Report 89, "An Evaluation of Using Job Performance Tests to
     Validate ASVAB Qualification Standards," by Milton H. Maier and
     Catherine M. Hiatt, Unclassified, May 1985